

# Online Research @ Cardiff

This is an Open Access document downloaded from ORCA, Cardiff University's institutional repository: <https://orca.cardiff.ac.uk/id/eprint/85189/>

This is the author's version of a work that was submitted to / accepted for publication.

Citation for final published version:

Cerutti, Federico ORCID: <https://orcid.org/0000-0003-0755-0358>, Kaplan, Lance M., Norman, Timothy J., Oren, Nir and Toniolo, Alice 2015. Subjective logic operators in trust assessment: an empirical study. Information Systems Frontiers 17 (4) , pp. 743-762. 10.1007/s10796-014-9522-5 file

Publishers page: <https://doi.org/10.1007/s10796-014-9522-5>  
<<https://doi.org/10.1007/s10796-014-9522-5>>

Please note:

Changes made as a result of publishing processes such as copy-editing, formatting and page numbers may not be reflected in this version. For the definitive version of this publication, please refer to the published source. You are advised to consult the publisher's version if you wish to cite this paper.

This version is being made available in accordance with publisher policies.

See

<http://orca.cf.ac.uk/policies.html> for usage policies. Copyright and moral rights for publications made available in ORCA are retained by the copyright holders.



# Subjective Logic Operators in Trust Assessment: An Empirical Study

Federico Cerutti · Lance M. Kaplan · Timothy J. Norman · Nir Oren · Alice Toniolo

Received: date / Accepted: date

**Abstract** Computational trust mechanisms aim to produce trust ratings from both direct and indirect information about agents' behaviour. Subjective Logic (SL) has been widely adopted as the core of such systems via its fusion and discount operators. In recent research we revisited the semantics of these operators to explore an alternative, geometric interpretation. In this paper we present principled desiderata for discounting and fusion operators in SL. Building upon this we present operators that satisfy these desirable properties, including a family of discount operators. We then show, through a rigorous empirical study, that specific, geometrically interpreted, operators significantly outperform standard SL operators in estimating ground truth. These novel operators offer real advantages for computational models of

trust and reputation, in which they may be employed without modifying other aspects of an existing system.

**Keywords** Trust and reputation · Information fusion · Uncertain reasoning

## 1 Introduction

Trust forms the backbone of human and artificial societies, improving robustness of interactions by restricting the actions of untrusted entities and mitigating the impact of untrusted information (Sensoy et al. 2013). Within the multi-agent systems community (Sabater and Sierra 2005), the problem of how to determine the degree of trustworthiness to assign to other agents is foundational for the notion of agency and for its defining relation of acting “on behalf of”. Trustworthiness is utilised when selecting partners for interactions; distrusted agents are less likely to be engaged, reducing their influence over the system.

Trust mechanisms aim to compute a level of trust based on direct and second-hand interactions between agents. The latter, commonly referred to as *reputational* information, is obtained from other agents which have interacted with the subject of the assessment. Aspects of such systems that have been examined include how to minimise the damage caused by collusion between agents (Haghpanah and des Jardins 2012), the nature of reputation information (Jøsang et al. 2012), and examining trust in specific contexts and agent interaction configurations (Burnett and Oren 2012).

In this paper we strengthen the analysis of an alternative to Jøsang's Subjective Logic (SL) discounting and combination operators (Jøsang 2001), which we have previously described in (Cerutti et al. 2013a). In particular, we enlarged the range of proposed discounting operators in order to provide a more comprehensive experimental evaluation. Instead of providing single operators we present a general approach,

Research was sponsored by US Army Research laboratory and the UK Ministry of Defence and was accomplished under Agreement Number W911NF-06-3-0001. The views and conclusions contained in this document are those of the authors and should not be interpreted as representing the official policies, either expressed or implied, of the US Army Research Laboratory, the U.S. Government, the UK Ministry of Defense, or the UK Government. The US and UK Governments are authorized to reproduce and distribute reprints for Government purposes notwithstanding any copyright notation hereon.

Corresponding author: F. Cerutti.

F. Cerutti · T. J. Norman · N. Oren · A. Toniolo  
University of Aberdeen  
School of Natural and Computing Science  
King's College  
AB24 3UE, Aberdeen, UK  
E-mail: f.cerutti@abdn.ac.uk  
E-mail: t.j.norman@abdn.ac.uk  
E-mail: n.oren@abdn.ac.uk  
E-mail: a.toniolo@abdn.ac.uk

L. Kaplan  
US Army Research Laboratory  
Adelphi  
MD, USA  
E-mail: lkapan@ieee.org

from which an entire family of operators can be derived that are proved to be compliant with specific desirable properties. From our analysis we can deduce some new interesting statistical properties of Jøsang's operators, as well as of the proposed operators. This evaluation methodology introduces two different metrics: the *expected value distance* from the ground truth, and the *geometric distance* from the ground truth. According to the former, our family of operators are shown to be almost equivalent to Jøsang's original operators, and significantly more accurate in one case. Using the latter metric, our operators compute reputation opinions closer to the ground truth than Jøsang's, with the exception of one case. Further, one of our proposed discounting operators outperforms the traditional SL operator on both metrics.

In the next section we present desiderata for discounting and fusion operators, grounded on how trust models such as SL are employed in practice. After a brief overview of Jøsang's SL in Section 3, we formalise the desiderata in Section 4 considering SL opinions, and show that they are not satisfied by existing SL operators. We describe our proposed operators in Section 5, and prove that they comply with the desirable properties presented. Then, in Section 6, we describe our experiment designed to conduct the comparative study among the operators, and, in Section 7, we evaluate our results to determine their significance. We summarise the conclusions that can be drawn from this study in Section 8. In order to improve the readability of the paper, proofs of the described results can be found in Appendix B (Appendix A discusses the mathematical foundations of the proposed operators).

## 2 Desiderata for Discounting and Fusion Operators

In this work we focus on trust relations where an agent referred to as the *truster* —  $X$  — “depends” on a *trustee* —  $Y_i$  — (Castelfranchi and Falcone 2010). As a concrete example, we examine the case where a trustee is responsible for providing some information to a truster, as exemplified by the following scenario.

**Example 1** *Let  $X$  be a military analyst who is collecting evidence in order to decide whether or not a specific area contains a certain type of weapon. In particular, he needs a datum  $m$  from two sensors  $Y_1$  and  $Y_2$ , each of which has a history of failure, thus affecting the trust  $X$  places in them. Here,  $X$  is the truster, and  $Y_1$  and  $Y_2$  the trustees.*

In such a scenario, the degree of trustworthiness of  $Y_1$  and  $Y_2$  is normally computed from historical data (Jøsang and McAnally 2004). Suppose that  $X$  asks  $Y_1$  and  $Y_2$  about  $m$ . Let us consider the case where  $Y$  informs that it believes that  $m$  holds (we write  $Y$  to indicate either  $Y_1$  or  $Y_2$ ).

**Example 1 (Continued) *Belief*.** *Suppose that  $Y$  answers that  $m$  holds with absolute certainty. However,  $Y$  has had a history of (random) failures, which means that  $X$  does not completely trust  $Y$ 's reports. In this situation, it seems reasonable for  $X$  to derive an opinion about the likelihood of  $m$  with an upper limit equivalent to its trust in  $Y$ .*

The above scenario gives us an intuition about a first desideratum concerning discounting opinions, viz.:

**desD<sub>1</sub>:** when the trustworthiness degree of a trustee  $Y$  is derived from historical data, if  $Y$  informs  $X$  that  $m$  holds with absolute certainty,  $X$  should believe  $m$  as much as it believes  $Y$ .

On the other hand, if  $Y$  informs  $X$  that it is uncertain about  $m$ , this should be directly reflected in  $X$ 's opinion about  $m$ .

**Example 1 (Continued) *Uncertainty*.** *Consider the case where  $Y$  informs  $X$  that it is unable to observe  $m$ . Here, there is complete uncertainty with regard to the degree of trustworthiness associated with  $Y$ 's reports about  $m$ .*

We can thus derive an additional desideratum:

**desD<sub>2</sub>:** if  $Y$  informs  $X$  that it is uncertain about  $m$ ,  $X$  is also uncertain about  $m$ .

In addition, we can identify an intermediate case, where it is known that the current situation negatively affects the degree of trustworthiness, and where an estimate of this effect can be determined. This is illustrated in the following scenario.

**Example 1 (Continued) *Intermediate evidence*.** *Suppose that  $Y$  reports that its opinion about  $m$  is not accurate, but there is some evidence in favour of  $m$ , and some evidence against it (with some degree of uncertainty)<sup>1</sup>. In this case  $X$  knows that the data received is somewhat accurate, and can therefore derive a degree of trustworthiness in the information regarding  $m$  received from  $Y$ .*

This illustrates another desideratum regarding the discounting of opinions, namely:

**desD<sub>3</sub>:** when the trustworthiness degree of a trustee  $Y$  is derived from historical data, if  $Y$  informs  $X$  that  $m$  holds with some degree of certainty,  $X$  should believe  $m$  no more than the trustworthiness degree of  $Y$ .

There are cases where the queried datum is not evidence about a datum, but rather an opinion about another agent.

<sup>1</sup> An example of this is GPS data, which is known to be inaccurate if you are using civilian equipment (Bisdikian et al. 2012).

**Example 1 (Continued) Reputation.** Suppose that sensor  $Y_1$  provides information about readings obtained from  $Y_2$ , an Uninterruptible Power Supply (UPS). Further suppose that  $X$  wants to query  $Y_2$  about its battery status, but  $X$  has, until now, only obtained its information directly from  $Y_1$ . Given a report from  $Y_2$ ,  $X$  can ask  $Y_1$  about  $Y_2$ 's degree of trustworthiness. According to our previous terminology, the message that  $Y_1$  sends to  $X$  is the subjective — to  $Y_1$  — trustworthiness degree of  $Y_2$ .

Finally, there are cases where  $X$  has to integrate different sources of information.

**Example 1 (Continued) Consensus.** Suppose that  $X$  queries both  $Y_1$  and  $Y_2$  about a datum  $m$ , and let us suppose that it receives different answers from the two sensors. In this situation,  $X$  will search for a consensus between these opinions, and will be biased towards the answer obtained from the more historically accurate sensor.

This illustrates a desideratum about fusing opinions:

**desF<sub>1</sub>:** in the case where  $X$  receives  $n$  opinions about  $m$  from  $Y_1, Y_2, \dots, Y_n$ , there must exist an operator informing which are  $X$ 's preferences among each  $Y_i$  (i.e. related to their degree of trustworthiness), and  $X$ 's opinion about  $m$  should be derived according to these preferences (i.e. giving more importance to the opinion received from the most preferred trustee).

### 3 Background

In the above scenario, we utilised the terms *trust*, *trustworthiness*, and *reputation* which often have different meanings among different pieces of research. It is beyond the scope of this paper to investigate these meanings; the interested reader is referred to (Castelfranchi and Falcone 2010) and (Urbano et al. 2013) for an overview. For the purpose of this paper we consider the notion of *trustworthiness* as the property of an agent — the *trustee* — that a *trustor* is connected with, where this property represents the willingness of the trustee to share information accurately with respect to the ground truth (we make a distinction between inaccuracy that is intentional or otherwise). Moreover, *reputation* is a property which represents the subjective view of an arbitrary trustee's trustworthiness that we obtained from an agent we can directly communicate with.

Following (Jøsang et al. 2007) we express both the *degree of trustworthiness* and the *degree of reputation* using SL. This formalism extends probability theory by expressing uncertainty about the probability values themselves, which makes it useful for representing trust degrees. We now proceed to provide a brief overview of SL mainly based on (Jøsang 2001).

Like Dempster-Shafer evidence theory (Dempster 1968; Shafer 1976), SL operates on a *frame of discernment*, denoted by  $\Theta$ . A frame of discernment contains the set of possible system states, only one of which represents the actual system state. These are referred to as atomic, or primitive, system states. The powerset of  $\Theta$ , denoted by  $2^\Theta$ , consists of all possible unions of primitive states. A non-primitive state may contain other states within it. These are referred to as substates of the state.

**Definition 1** Given a frame of discernment  $\Theta$ , we can associate a belief mass assignment  $m_\Theta(x)$  with each substate  $x \in 2^\Theta$  such that

1.  $m_\Theta(x) \geq 0$
2.  $m_\Theta(\emptyset) = 0$
3.  $\sum_{x \in 2^\Theta} m_\Theta(x) = 1$

For a substate  $x$ ,  $m_\Theta(x)$  is its *belief mass*.

Belief mass is an unwieldy concept to work with. When we speak of belief in a certain state, we refer not only to the belief mass in the state, but also to the belief masses of the state's substates. Similarly, when we speak about disbelief, that is, the total belief that a state is not true, we need to take substates into account. Finally, SL also introduces the concept of uncertainty, that is, the amount of belief that might be in a superstate or a partially overlapping state. These concepts can be formalised as follows.

**Definition 2** Given a frame of discernment  $\Theta$  and a belief mass assignment  $m_\Theta$  on  $\Theta$ , we define the belief function for a state  $x$  as

$$b(x) = \sum_{y \subseteq x} m_\Theta(y) \text{ where } x, y \in 2^\Theta$$

The disbelief function as

$$d(x) = \sum_{y \cap x = \emptyset} m_\Theta(y) \text{ where } x, y \in 2^\Theta$$

And the uncertainty function as

$$u(x) = \sum_{\substack{y \cap x \neq \emptyset \\ y \not\subseteq x}} m_\Theta(y) \text{ where } x, y \in 2^\Theta$$

These functions have two important properties. First, they all range between zero and one. Second, they always sum to one, meaning that it is possible to deduce the value of one function given the other two.

Boolean logic operators have SL equivalents. It makes sense to use these equivalent operators in frames of discernment containing a state and (some form of) the state's negation. A *focused frame of discernment* is a binary frame of discernment containing a state and its complement.

**Definition 3** Given  $x \in 2^\Theta$ , the frame of discernment denoted by  $\tilde{\Theta}^x$ , which contains two atomic states,  $x$  and  $\neg x$ , where  $\neg x$  is the complement of  $x$  in  $\Theta$ , is the focused frame of discernment with focus on  $x$ .

Let  $\tilde{\Theta}^x$  be the focused frame of discernment with focus on  $x$  of  $\Theta$ . Given a belief mass assignment  $m_\Theta$  and the belief, disbelief and uncertainty functions for  $x$  ( $b(x)$ ,  $d(x)$  and  $u(x)$  respectively), the focused belief mass assignment,  $m_{\tilde{\Theta}^x}$  on  $\tilde{\Theta}^x$  is defined as

$$\begin{aligned} m_{\tilde{\Theta}^x}(x) &= b(x) \\ m_{\tilde{\Theta}^x}(\neg x) &= d(x) \\ m_{\tilde{\Theta}^x}(\tilde{\Theta}^x) &= u(x) \end{aligned}$$

The focused relative atomicity of  $x$  (which approximates the role of a prior probability distribution within probability theory, weighting the likelihood of some outcomes over others) is defined as

$$a_{\tilde{\Theta}^x}(x/\Theta) = [E(x) - b(x)]/u(x)$$

where  $E(x)$  represents the *expected value* of  $x$ .

For convenience, the focused relative atomicity of  $x$  is often abbreviated  $a_{\tilde{\Theta}^x}(x)$  or  $a(x)$ .

An opinion consists of the belief, disbelief, uncertainty and relative atomicity as computed over a focused frame of discernment.

**Definition 4** Given a focused frame of discernment  $\Theta$  containing  $x$  and its complement  $\neg x$ , and assuming a belief mass assignment  $m_\Theta$  with belief, disbelief, uncertainty and relative atomicity functions on  $x$  in  $\Theta$  of  $b(x)$ ,  $d(x)$ ,  $u(x)$  and  $a(x)$ , we define an *opinion* over  $x$ , written  $\omega_x$  as

$$\omega_x \equiv \langle b(x), d(x), u(x), a(x) \rangle$$

For compactness, Jøsang also denotes the various functions as  $b_x, d_x, u_x$  and  $a_x$  in place, and we will follow this notation. Furthermore, given a fixed  $a_x$ , an opinion  $\omega$  can be denoted as a  $\langle b_x, d_x, u_x \rangle$  triple.

Given opinions about two propositions from different frames of discernment, it is possible to combine them in various ways using SL's various operators, as detailed in (Jøsang 2001; Jøsang and McAnally 2004; Jøsang et al. 2005, 2006; McAnally and Jøsang 2004). In this work we concentrate on Jøsang's *discount* and *fusion* operators, which we review next.

**Definition 5 (Def. 5, (Jøsang et al. 2006))** Let  $A, B$  be two agents where  $A$ 's opinion about  $B$  is expressed as  $\omega_B^A = \langle b_B^A, d_B^A, u_B^A, a_B^A \rangle$  and let  $x$  be a proposition where  $B$ 's opinion about  $x$  is acquired by  $A$  as the opinion  $\omega_x^B = \langle b_x^B, d_x^B, u_x^B, a_x^B \rangle$ . Let  $\omega_x^{A:B} = \langle b_x^{A:B}, d_x^{A:B}, u_x^{A:B}, a_x^{A:B} \rangle$  be the opinion such that:

$$\begin{cases} b_x^{A:B} = b_B^A b_x^B \\ d_x^{A:B} = b_B^A d_x^B \\ u_x^{A:B} = d_B^A + u_B^A + b_B^A u_x^B \\ a_x^{A:B} = a_x^B \end{cases}$$

then  $\omega_x^{A:B}$  is called the *uncertainty favouring discounted opinion* of  $A$ . By using the symbol  $\otimes$  to designate this operation, we get  $\omega_x^{A:B} = \omega_B^A \otimes \omega_x^B$ .

**Definition 6 (Thm. 1, (Jøsang et al. 2006))** Let  $\omega_x^A = \langle b_x^A, d_x^A, u_x^A, a_x^A \rangle$  and  $\omega_x^B = \langle b_x^B, d_x^B, u_x^B, a_x^B \rangle$  be trust in  $x$  from  $A$  and  $B$  respectively. The opinion  $\omega_x^{A \diamond B} = \langle b_x^{A \diamond B}, d_x^{A \diamond B}, u_x^{A \diamond B}, a_x^{A \diamond B} \rangle$  is then called the consensus between  $\omega_x^A$  and  $\omega_x^B$ , denoting the trust that an imaginary agent  $[A, B]$  would have in  $x$ , as if that agent represented both  $A$  and  $B$ . In case of Bayesian (totally certain) opinions, their relative weight can be defined as  $\gamma^{A/B} = \lim (u_x^B / u_x^A)$ .

Case I:  $u_x^A + u_x^B - u_x^A u_x^B \neq 0$

$$\begin{cases} b_x^{A \diamond B} = \frac{b_x^A u_x^B + b_x^B u_x^A}{u_x^A + u_x^B - u_x^A u_x^B} \\ d_x^{A \diamond B} = \frac{d_x^A u_x^B + d_x^B u_x^A}{u_x^A + u_x^B - u_x^A u_x^B} \\ u_x^{A \diamond B} = \frac{u_x^A u_x^B}{u_x^A + u_x^B - u_x^A u_x^B} \\ a_x^{A \diamond B} = \frac{a_x^A u_x^B + a_x^B u_x^A - (a_x^A + a_x^B) u_x^A u_x^B}{u_x^A + u_x^B - 2 u_x^A u_x^B} \end{cases}$$

Case II:  $u_x^A + u_x^B - u_x^A u_x^B = 0$

$$\begin{cases} b_x^{A \diamond B} = \frac{(\gamma^{A/B} b_x^A + b_x^B)}{(\gamma^{A/B} + 1)} \\ d_x^{A \diamond B} = \frac{(\gamma^{A/B} d_x^A + d_x^B)}{(\gamma^{A/B} + 1)} \\ u_x^{A \diamond B} = 0 \\ a_x^{A \diamond B} = \frac{(\gamma^{A/B} a_x^A + a_x^B)}{(\gamma^{A/B} + 1)} \end{cases}$$

By using the symbol  $\oplus$  to designate this operator, we can write  $\omega_x^{A \diamond B} = \omega_x^A \oplus \omega_x^B$ .

## 4 Core Properties and Requirements

In our scenario, agent  $X$  has to determine the trustworthiness degree associated with  $m$  received from  $Y_i$ .  $X$  will consider three elements in reaching a decision:

1. **trustworthiness:**  $X$  has an opinion  $T_i$  concerning the degree of trustworthiness of  $Y_i$ ;
2. **certainty:**  $Y_i$  communicates that  $m$  holds with a degree of certainty  $C_i$ ;
3. **combination:**  $X$  has to combine  $T_1, \dots, T_n$  with (resp.)  $C_1, \dots, C_n$  in order to achieve an ultimate opinion  $W = \Gamma(W_1, \dots, W_n)$  on  $m$ , where  $\forall i W_i = T_i \circ C_i$ , i.e.  $W_i$  is the result of a combination of opinion  $T_i$  with opinion  $C_i$ , and each opinion  $W_i = T_i \circ C_i$  has associated a weight  $K_i = f(T_i)$  for some function  $f(\cdot)$ .

In particular, the three desiderata for discounting opinions (desD<sub>1</sub>, desD<sub>2</sub>, desD<sub>3</sub>), give rise to the following *requirements for discounting*:

- rd<sub>1</sub>** if  $C_i$  is pure belief, then  $W_i = T_i$ ;  
**rd<sub>2</sub>** if  $C_i$  is completely uncertain, then  $W_i = C_i$  (i.e.  $\langle 0, 0, 1 \rangle$ );  
**rd<sub>3</sub>** the degree of belief of  $W_i$  is always less than or equal to the degree of belief of  $T_i$ ;

and the desideratum about fusing opinion (**desF<sub>1</sub>**) gives rise to the following *requirements for fusion*:

- rf<sub>1</sub>** if  $\forall i, j \ K_i = K_j$ , then  $W = \Gamma(W_1, \dots, W_n)$  is the opinion resulting from the average of  $W_1, \dots, W_n$ ;  
**rf<sub>2</sub>** if  $\exists i$  s.t.  $K_i = 0$ , then  $\Gamma(W_1, \dots, W_n) = \Gamma(W_1, \dots, W_{i-1}, W_{i+1}, \dots, W_n)$ .

While there is a direct correspondence between the desideratum for discounting opinion (**desD<sub>1</sub>**, **desD<sub>2</sub>**, **desD<sub>3</sub>**) and the requirements for discounting (**rd<sub>1</sub>**, **rd<sub>2</sub>**, **rd<sub>3</sub>**), the desideratum about fusing opinion **desF<sub>1</sub>** gives rise to two (loose) requirements, **rf<sub>1</sub>** and **rf<sub>2</sub>**. Note that the combination requirements describe the same “prudent” behaviour as was presented in Example 1, in particular in the “belief” scenario. Indeed even if  $X$  is highly confident in a specific context, this confidence cannot increase the trust degree over the base trustworthiness degree.

Following (McAnally and Jøsang 2004; Jøsang et al. 2006), we utilise SL to instantiate trustworthiness and confidence, and seek to compute their combination through SL operations<sup>2</sup>. In doing so, we must therefore consider the following inputs and requirements:

1.  $\mathbb{O} = \{\langle b, d, u \rangle \in \mathbb{R}^3 \mid 0 \leq b \leq 1 \text{ and } 0 \leq d \leq 1 \text{ and } 0 \leq u \leq 1 \text{ and } b + d + u = 1\}$ ;
2.  $T_i = \langle b_{T_i}, d_{T_i}, u_{T_i} \rangle$  derived by statistical observations (e.g. (McAnally and Jøsang 2004));
3.  $C_i = \langle b_{C_i}, d_{C_i}, u_{C_i} \rangle$  is the opinion received from  $Y_i$  about  $m$ ;
4.  $\circ : \mathbb{O} \times \mathbb{O} \mapsto \mathbb{O}$ ;
5.  $\Gamma : \mathbb{O}^n \mapsto \mathbb{O}$ ;
6.  $W_i = \begin{cases} T_i & \text{if } C_i = \langle 1, 0, 0 \rangle \\ C_i & \text{if } C_i = \langle 0, 0, 1 \rangle \end{cases}$   
further requiring that  $b_{W_i} \leq b_{T_i}$ ;
7. given  $W = \Gamma(W_1, \dots, W_n)$ , if  $\forall i, j \in \{1, \dots, n\} K_i = K_j$ , then
  - $b_W = \frac{1}{n} \sum_{i=1}^n b_{W_i}$
  - $d_W = \frac{1}{n} \sum_{i=1}^n d_{W_i}$
  - $u_W = \frac{1}{n} \sum_{i=1}^n u_{W_i}$
8. if  $\exists i$  s.t.  $K_i = 0$ , then  $\Gamma(W_1, \dots, W_n) = \Gamma(W_1, \dots, W_{i-1}, W_{i+1}, \dots, W_n)$ .

Although any function  $f(\cdot)$  can be used for deriving  $K_i$ , hereafter we will consider  $T_i$ ’s expected value, i.e.  $K_i = b_{T_i} + \frac{u_{T_i}}{2}$ .

<sup>2</sup> Hereafter each opinion will have a fixed relative atomicity of  $\frac{1}{2}$ . This assumption will be relaxed in future work.

Since 1–5 above are inputs, we concentrate on the constraints expressed by 6, 7 and 8, which require us to consider the problem of how to combine the degree of trustworthiness with the degree of certainty. Existing work, such as (McAnally and Jøsang 2004; Castelfranchi and Falcone 2010; Urbano et al. 2013), concentrate on computing  $T$ .

We begin by noting — as illustrated in Table 1 — that no set of operators provided by SL (Jøsang 2001; Jøsang and McAnally 2004; Jøsang et al. 2005, 2006; McAnally and Jøsang 2004; Jøsang 2008) satisfies the combination requirements previously described.

Operator	Requirement satisfied?				
	Discount req.			Fusion req.	
	<b>rd<sub>1</sub></b>	<b>rd<sub>2</sub></b>	<b>rd<sub>3</sub></b>	<b>rf<sub>1</sub></b>	<b>rf<sub>2</sub></b>
Addition (+)	No	No	No	No	No
Subtraction (−)	No	No	Yes	No	No
Multiplication (·)	No	No	No	No	No
Division (/)	No	No	No	No	No
Comultiplication ( $\sqcup$ )	No	No	No	No	No
Codivision ( $\sqcap$ )	No	No	No	No	No
Discounting ( $\otimes$ )	No	Yes	Yes	No	No
Cumulative fusion ( $\oplus$ )	No	No	No	No	No
Averaging fusion ( $\oplus$ )	No	No	No	Yes	No
Cumulative unfusion ( $\ominus$ )	No	No	No	No	No
Averaging unfusion ( $\ominus$ )	No	No	No	No	No

Table 1: Jøsang operators and the satisfaction of the five combination requirements

In the next section, we describe our proposals for the discount —  $\circ$  — and consensus —  $\Gamma$  — operators in order to satisfy the above five requirements.

## 5 The Operators

### 5.1 A Naïve Discount Operator

As suggested by an anonymous referee of a preliminary version of (Cerutti et al. 2013b), a very naïve operator satisfying the requirements **rd<sub>1</sub>**, **rd<sub>2</sub>**, and **rd<sub>3</sub>** is the following.

**Definition 7** Given the two opinions  $T = \langle b_T, d_T, u_T \rangle$  and  $C = \langle b_C, d_C, u_C \rangle$ , the naïve-discount of  $C$  by  $T$  is  $W = T \circ_n C$ , where:

- $b_W = b_C \cdot b_T$ ;
- $d_W = b_C \cdot d_T + d_C$ ;
- $u_W = b_C \cdot u_T + u_C$ .

The following proposition shows that the naïve discount operator fulfils the first three requirements.

**Proposition 1** Given the two opinions  $T = \langle b_T, d_T, u_T \rangle$  and  $C = \langle b_C, d_C, u_C \rangle$ , and  $W = T \circ_n C$ , the naïve-discount of  $C$  by  $T$ , then:





$$\begin{aligned}
- \gamma_T &= \frac{\pi}{3} - \beta_T \\
- \delta_T &= \begin{cases} 0 & \text{if } u_T = 1 \\ \arcsin\left(\frac{b_T}{|\overrightarrow{TU}|}\right) & \text{otherwise} \end{cases} \\
- \epsilon_T &= \pi - \gamma_T - \delta_T \\
- |\overrightarrow{TC'}| &= \frac{|\overrightarrow{BC}|}{|\overrightarrow{BM_C}|} |\overrightarrow{TM_{C'}}| = \\
&= r_C |\overrightarrow{TM_{C'}}|
\end{aligned}$$

with  $r_C = \frac{|\overrightarrow{BC}|}{|\overrightarrow{BM_C}|} = 1 - b_C$ , and

$$|\overrightarrow{TM_{C'}}| = \begin{cases} 2b_T & \text{if } \alpha_{C'} = \frac{\pi}{2} \\ \frac{2}{\sqrt{3}} u_T & \text{if } \alpha_{C'} = -\frac{\pi}{3} \\ \frac{2}{\sqrt{3}} (1 - u_T) & \text{if } \alpha_{C'} = \frac{2}{3}\pi \\ \frac{2\sqrt{\tan^2(\alpha_{C'}) + 1}}{|\tan(\alpha_{C'}) + \sqrt{3}|} b_T & \text{otherwise} \end{cases}$$

We can now show that this family of operators satisfies the requirements **rd**<sub>1</sub>, **rd**<sub>2</sub>, and **rd**<sub>3</sub>.

**Theorem 1** Given  $T = \langle b_T, d_T, u_T \rangle$ ,  $C = \langle b_C, d_C, u_C \rangle$ , and  $W = T \circ_{\alpha_{C'}} C$ , then:

- i.  $W = \langle b_W, d_W, u_W \rangle$  is an opinion;
- ii. if  $C = \langle 1, 0, 0 \rangle$ , then  $W = T$ ;
- iii. if  $C = \langle 0, 0, 1 \rangle$ , then  $W = C$ ;
- iv.  $b_W \leq b_T$ .

In order to empirically examine the difference among members of this family of operators, let us define the following three graphical discount operators. The first,  $\circ_1$ , considers the widest range of  $\alpha_{C'}$  possible, projecting  $C$  in the triangle  $\triangle DTU$ . This operator is a “geometrical counterpart” to the naïve operator — projecting  $C$  in  $\triangle DTU$  could result in an opinion with less uncertainty than  $T$ . The second operator,  $\circ_2$ , raises the uncertainty of the discount opinion projecting  $C$  in the triangle  $\triangle RTU$ . The third operator,  $\circ_3$ , considers the triangle  $\triangle STU$ , where  $S$  is the intersection of the line  $DU$  with the bisector of the angle  $\epsilon_T$  (this point is not shown in Fig. 1).

**Definition 10** Given the two opinions  $T = \langle b_T, d_T, u_T \rangle$  and  $C = \langle b_C, d_C, u_C \rangle$ :

- $\circ_1$  is  $\overline{\circ_{\alpha_{C'}}$  with  $\alpha_{C'} = \frac{\alpha_C \epsilon_T}{\frac{\pi}{3}} - \beta_T$ ;
- $\circ_2$  is  $\overline{\circ_{\alpha_{C'}}$  with  $\alpha_{C'} = \frac{\alpha_C (\epsilon_T - \beta_T)}{\frac{\pi}{3}}$ ;
- $\circ_3$  is  $\overline{\circ_{\alpha_{C'}}$  with  $\alpha_{C'} = \frac{\alpha_C \frac{\epsilon_T}{2}}{\frac{\pi}{3}} + \frac{\epsilon_T}{2} - \beta_T$ .

### 5.3 The Graphical Fusion Operator

In (Cerutti et al. 2013a) we introduced a fusion operator which satisfies requirements **rf**<sub>1</sub> and **rf**<sub>2</sub>. Let us suppose we have  $n$  opinions  $W_1, W_2, \dots, W_n$  derived using an operator  $\circ$  s.t.  $\forall i \in \{1, \dots, n\}, W_i = T_i \circ C_i$ . Intuitively, the fused opinion  $\Gamma(W_1, W_2, \dots, W_n)$  should be the “balanced” centroid of the polygon determined by the  $n$  opinions.

**Definition 11** Given the opinions  $T_1, T_2, \dots, T_n, C_1, C_2, \dots, C_n, W_1, W_2, \dots, W_n$  such that  $\forall i \in \{1 \dots n\}, W_i = T_i \circ C_i$ , the opinion resulting from the fusion of opinions  $W_1, W_2, \dots, W_n$  is

$$\langle b_{\Gamma_1(W_1, \dots, W_n)}, d_{\Gamma_1(W_1, \dots, W_n)}, u_{\Gamma_1(W_1, \dots, W_n)} \rangle$$

where:

$$\begin{aligned}
- b_{\Gamma_1(W_1, \dots, W_n)} &= \frac{1}{\sum_{i=1}^n K_i} \left( \sum_{i=1}^n K_i b_{W_i} \right) \\
- d_{\Gamma_1(W_1, \dots, W_n)} &= \frac{1}{\sum_{i=1}^n K_i} \left( \sum_{i=1}^n K_i d_{W_i} \right) \\
- u_{\Gamma_1(W_1, \dots, W_n)} &= \frac{1}{\sum_{i=1}^n K_i} \left( \sum_{i=1}^n K_i u_{W_i} \right)
\end{aligned}$$

This definition of  $\Gamma_1$  satisfies the requirements **rf**<sub>1</sub> and **rf**<sub>2</sub>. Moreover, the following proposition shows that  $\Gamma_1(W_1, \dots, W_n)$  is an opinion, and its Cartesian representation is the balanced centroid of the polygon identified by the points  $W_1, \dots, W_n$ .

**Proposition 3** Given the opinions  $T_1, T_2, \dots, T_n, C_1, C_2, \dots, C_n, W_1, W_2, \dots, W_n$  s.t.  $\forall i \in \{1 \dots n\}, W_i = T_i \circ C_i$ , and  $\langle b_{\Gamma_1(W_1, \dots, W_n)}, d_{\Gamma_1(W_1, \dots, W_n)}, u_{\Gamma_1(W_1, \dots, W_n)} \rangle$  the opinion resulting from the fusion of opinions  $W_1, W_2, \dots, W_n$ , then:

- i.  $\langle b_{\Gamma_1(W_1, \dots, W_n)}, d_{\Gamma_1(W_1, \dots, W_n)}, u_{\Gamma_1(W_1, \dots, W_n)} \rangle$  is an opinion
- ii.  $x_{\Gamma_1(W_1, \dots, W_n)} = \frac{1}{\sum_{i=1}^n K_i} \left( \sum_{i=1}^n K_i x_{W_i} \right)$
- iii.  $y_{\Gamma_1(W_1, \dots, W_n)} = \frac{1}{\sum_{i=1}^n K_i} \left( \sum_{i=1}^n K_i y_{W_i} \right)$

## 6 Experimental Evaluation

In Section 2 we discussed the desiderata for the discounting and fusion operators. In Section 4, we used these to obtain requirements **rd**<sub>1</sub>, **rd**<sub>2</sub>, **rd**<sub>3</sub>, **rf**<sub>1</sub>, **rf**<sub>2</sub>, which our proposed operators have been shown to fulfil.

In this section we focus on the design of an experiment aimed at evaluating, from an empirical point of view, the significance and usefulness of the desiderata and the requirements we identified for discounting and fusion operators. To



this end, we consider an experiment where an explorer agent has to explore a network in order to determine the trustworthiness degree of other members: some of them are directly connected to the explorer, while for others it has to discount reputational information, which is provided in the form of SL opinions. The evaluation is based upon the distance between the derived opinion and the ground truth, i.e. the probability that each agent operates in a trustworthy manner. This likelihood is instantiated with the system, and is not known by anyone else in the network. This is just one of the possible scenarios where the discounting and fusion operators can be applied: other empirical evaluation is left as avenues of future research.

The experimental setup follows the one described in (Cerutti et al. 2013a). In this experiment each agent can communicate with all the other agents in the network. In order to randomly generate these networks, we consider a variable  $P^L \in [0, 1]$  representing the probability that an agent is connected<sup>6</sup> to another agent (we exclude self-connections). Note that we do not constrain connections to be bidirectional<sup>7</sup>. In this experiment, we set  $P^L$  to lie between 0.05 and 0.25 with increments of 0.05. For each value of  $P^L$ , we execute the procedure described in the following sections.

### 6.1 Trust System Construction

We build a set of 50 agents  $\mathcal{A} = \{a_1, \dots, a_{50}\}$ : each agent  $a_x$  is characterised by a knowledge base  $\mathcal{KB}_{a_x}$  and by the probability of responding truthfully to another agent's query, namely  $P_{a_x}^T \in [0 \dots 1]$ . For each agent  $a_x$ , we randomly choose  $P_{a_x}^T$ .

We also require that  $(\Omega = \top) \in \mathcal{KB}_{a_x}$ . In other words, all the agents share the same information  $(\Omega = \top)$  to be read “ $a_x$  knows that  $\Omega$  is  $\top$ ”.

For each agent  $a_x$ , we determine if it can communicate with  $a_y \neq a_x$  according to  $P^L$ : if  $a_y$  is connected to  $a_x$ , then we say that  $a_y$  is a connection of  $a_x$  ( $a_y \in N_{a_x}$ ).

Following the construction of the system, the experiment proceeds through two distinct phases, namely bootstrapping and exploration.

### 6.2 Phase I: Bootstrapping

The bootstrapping phase is similar to that described in (Ismael and Jøsang 2002), where a beta distribution is used for

analysing repetitive experiments and deriving a SL opinion. In this experiment, each agent  $a_x$  asks  $\#_B$  times each of its connections  $a_y \in N_{a_x}$  its opinion about  $\Omega$ . Each time,  $a_y$  provides a possible false answer, based only on the probability  $P_{a_y}^T$  (the communications are stateless): the two possible answers of  $a_y$  are, of course,  $\Omega = \top$  and  $\Omega = \perp$ .

Agent  $a_x$  counts the number of exchanges where  $a_y$  answered truthfully ( $\#_\top$ ) and when it lied ( $\#_\perp$ ). Clearly,  $\#_B = \#_\top + \#_\perp$ . Using this evidence,  $a_x$  can form an opinion on  $a_y$ 's trustworthiness

$$O_{a_x}^{a_y} = \langle \frac{\#_\top}{\#_B + 2}, \frac{\#_\perp}{\#_B + 2}, \frac{2}{\#_B + 2} \rangle$$

which should be close (according to the definition of distance given in Def. 12) to the “ideal” (“real”) opinion the (omniscient) experimenter has on  $a_y$ , viz.

$$O_{Exp}^{a_y} = \langle P_{a_y}^T, (1 - P_{a_y}^T), 0 \rangle$$

Therefore, during the bootstrapping phase, each agent  $a_y \in N_{a_x}$ ,  $a_x$  records its opinion of  $a_y$  in its knowledge base.

In the experiment described in (Cerutti et al. 2013a),  $\#_B$  varies between 25 and 250: however, as we noted in that paper, this variation did not alter the results of the experiment. Therefore, in this paper we collected data for ten different values of  $\#_B$  varying it between 2 and 29 with a step size of 3.

### 6.3 Phase II: Exploration

After each agent has enriched its knowledge base with opinions of the trustworthiness of its connections, an “explorer”  $a_S \in \mathcal{A}$  is randomly selected. The task of this explorer is to determine the trustworthiness of each agent in the network. The explorer,  $a_S$ , acquires information about the network by asking its connections “Who are your connections?”. Each agent  $a_y \in N_{a_S}$  answers this question according to  $P_{a_y}^T$ , which means that their answers are: for each  $a_y \in N_{a_x}$   $Connections_{a_y} \subseteq N_{a_y}$  (clearly if  $P_{a_y}^T = 1$ , then  $Connections_{a_y} = N_{a_y}$ ).

Agent  $a_S$  collects all the answers and creates a set of tuples associating agents that the explorer does not directly know, and all the agents that have revealed that they have connections to that agent, viz.:

$$\mathcal{M} = \{ \langle a_z, \{a_{y_1}, \dots, a_{y_n}\} \rangle \mid \forall i \in [1 \dots n] \ a_{y_i} \in N_{a_S} \text{ and } a_z \in \bigcap_{i=1}^n N_{a_{y_i}} \}$$

Then, for each pair of  $\mathcal{M}$ ,  $\langle a_z, \{a_{y_1}, \dots, a_{y_n}\} \rangle$ , such that  $a_z \notin N_{a_S} \cup \{a_S\}$  (i.e. for each agent it is not directly

<sup>6</sup> The term “connected” here can have different names in different contexts, like “friend” in Facebook, or “follower” in Twitter.

<sup>7</sup> Although this may seem counter-intuitive, it partially captures real-world social media. For instance, Twitter messages are public, therefore we do not know who will read our messages. The same applies with slight modifications to Google+, and, of course, to blogging activities in general.

connected to),  $a_S$  asks each  $a_{y_i}$  (i.e. those that are connected to that agent) about  $O_{a_{y_i}}^{a_z}$  (i.e. their opinion of that agent). Agent  $a_{y_i}$  answers according to  $P_{a_{y_i}}^T$  either  $O_{a_{y_i}}^{a_z}$  or  $\langle b_R, d_R, u_R \rangle$  where  $\langle b_R, d_R, u_R \rangle$  is a SL opinion computed randomly such that  $\langle b_R, d_R, u_R \rangle \neq O_{a_{y_i}}^{a_z}$ . Since  $a_S$  cannot determine whether the answer is true or not, we abuse notation, by associating  $O_{a_{y_i}}^{a_z}$  with the answer  $a_S$  received from  $a_{y_i}$  to the question “What is your opinion about  $a_z$ ?”.

Subsequently,  $a_S$  computes  $O_{a_S|J}^{a_z} = (O_{a_S}^{a_{y_1}} \otimes O_{a_{y_1}}^{a_z}) \oplus \dots \oplus (O_{a_S}^{a_{y_n}} \otimes O_{a_{y_n}}^{a_z})$  (viz. the fusion of the discounted opinions on  $a_z$  of its connections using Jøsang’s operators), and, for each operator defined in Def. 7 and Def. 10 (i.e.  $\circ \in \{\circ_n, \circ_1, \circ_2, \circ_3\}$ )  $O_{a_S|\circ}^{a_z} = \Gamma_1((O_{a_S}^{a_{y_1}} \circ O_{a_{y_1}}^{a_z}), \dots, (O_{a_S}^{a_{y_n}} \circ O_{a_{y_n}}^{a_z}))$  (viz. the fusion of the discounted opinions on  $a_z$  of its connections using the naïve operator, and the members of the graphical operator family introduced in Def. 9 with the fusion operator of Def. 11). Since we considered only one fusion operator, viz.  $\Gamma_1$ , we will write it without a subscript to improve readability. Moreover, since we want to evaluate the proposed operators and compare them to Jøsang’s ones, each exploration has been performed with Jøsang’s operators and with only one of the proposed operators. We, therefore, explore the same network four times with Jøsang’s operator. However, doing so guarantees that the evaluation of each operator is independent from other evaluations.

Finally, each agent  $a_z$  is added to the list of the connections of  $a_S$  and the process starts again by setting  $\mathcal{M} = \emptyset$  and querying each member of the connections until, in two subsequent interactions, no further agents are added to  $a_S$ ’s connections. This exploration process, therefore enables  $a_S$  to form a picture of those agents in the network, which are not directly connected to  $a_S$ , by the opinions of other agents. Note that the results obtained through iteration  $j$  of the exploration phase serves to bootstrap iteration  $j + 1$ .

#### 6.4 Computing the Distances.

For each agent  $a_z \in \mathcal{A} \setminus \{a_S\}$ , for each  $\circ \in \{\circ_n, \circ_1, \circ_2, \circ_3\}$ , we compute the distance between the two derived opinions  $O_{a_S|J}^{a_z}$  and  $O_{a_S|\circ}^{a_z}$ , and the “ideal” opinion  $O_{Exp}^{a_z}$ . For this purpose, we consider two notions of distance, namely a *geometrical distance*, and a *distance of expected values*.

The *geometric distance* between two opinions  $\langle b_{O_1}, d_{O_1}, u_{O_1} \rangle$  and  $\langle b_{O_2}, d_{O_2}, u_{O_2} \rangle$  is the Euclidean distance between the two points in the  $\mathbb{O}$  space.

**Definition 12** Given two opinions  $O_1 = \langle b_{O_1}, d_{O_1}, u_{O_1} \rangle$  and  $O_2 = \langle b_{O_2}, d_{O_2}, u_{O_2} \rangle$ , the *geometric distance* between  $O_1$  and  $O_2$  is

$$d_G(O_1, O_2) = \sqrt{(b_{O_2} - b_{O_1})^2 + (d_{O_2} - d_{O_1})^2 + (u_{O_2} - u_{O_1})^2}$$

Another interesting distance measure is the difference between the *expected values* of subjective logic opinions.

Let us recall that the expected value for a subjective logic opinion  $\langle b_X, d_X, u_X, a_X \rangle$  is  $b_X + u_X \cdot a_X$ . Given that we assume a fixed base rate of  $\frac{1}{2}$ , the expected value simplifies to  $b_X + \frac{u_X}{2}$ . Given the expected value of two opinions, we can easily compute the distance between them.

**Definition 13** Given two opinions  $O_1 = \langle b_{O_1}, d_{O_1}, u_{O_1} \rangle$  and  $O_2 = \langle b_{O_2}, d_{O_2}, u_{O_2} \rangle$ , the *expected value distance* between  $O_1$  and  $O_2$  is

$$d_E(O_1, O_2) = \left| \left( b_{O_1} + \frac{u_{O_1}}{2} \right) - \left( b_{O_2} + \frac{u_{O_2}}{2} \right) \right|$$

In other words,  $\forall a_z \in \mathcal{A} \setminus \{a_S\}$ ,  $d_G(O_{a_S|J}^{a_z}, O_{Exp}^{a_z})$  (reps.  $d_E(O_{a_S|J}^{a_z}, O_{Exp}^{a_z})$ ) is the geometric distance (resp. expected value distance) between the derived opinion using Jøsang’s operators and the “ideal” one, and,  $\forall \circ \in \{\circ_n, \circ_1, \circ_2, \circ_3\}$ ,  $d_G(O_{a_S|\circ}^{a_z}, O_{Exp}^{a_z})$  (resp.  $d_E(O_{a_S|\circ}^{a_z}, O_{Exp}^{a_z})$ ) is the geometric distance (resp. expected value distance) between the derived opinion using either the naïve operators or the operators of the family of graphical discount operators (Def. 9) and the “ideal” one (abbrev.  $d_G^\circ$ , resp.  $d_E^\circ$ ).

Finally, for each  $a_z \in \mathcal{A} \setminus \{a_S\}$ ,  $\forall \circ \in \{\circ_n, \circ_1, \circ_2, \circ_3\}$ , we compare the two computed distances obtaining the following scalar comparison values:

$$r_G(a_z) = \log \frac{d_G(O_{a_S|J}^{a_z}, O_{Exp}^{a_z})}{d_G(O_{a_S|\circ}^{a_z}, O_{Exp}^{a_z})}$$

and

$$r_E(a_z) = \log \frac{d_E(O_{a_S|J}^{a_z}, O_{Exp}^{a_z})}{d_E(O_{a_S|\circ}^{a_z}, O_{Exp}^{a_z})}$$

To strengthen the significance of the results, we mainly concentrate on averages, and thus  $a_S$  explores the network  $|\mathcal{A}|/2 = 25$  times; we write  $r_G(a_z)$  (resp.  $r_E(a_z)$ ) to denote the average of the 25 computed logarithmic ratios using the geometric distance (resp. expected value distance). Moreover  $\overline{r_G(a_z)} = \text{average}_{a_z \in \mathcal{A} \setminus \{a_S\}} r(a_z)$  (resp.  $\overline{r_E(a_z)} = \text{average}_{a_z \in \mathcal{A} \setminus \{a_S\}} r(a_z)$ ) is the average of the comparison value over the whole set of agents using the geometric distance (resp. expected value distance).

## 7 Analysis of Experimental Results

To ensure that the outcomes are not biased by the random generator, we run the same experiment ten times. Each run follows the steps described in Sect. 6, and thus for each value of  $P^L$ , 10 networks have been randomly generated. Moreover, since each agent can lie, each generated network has been explored 25 times. Therefore, for each run, for each value of  $P^L$ , 250 explorations over 10 different networks

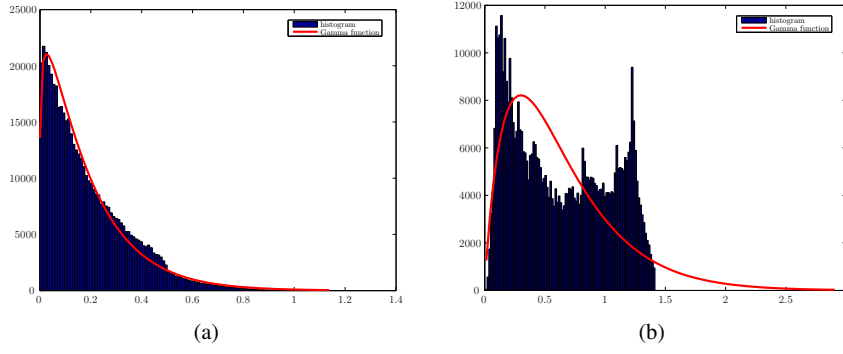


Fig. 2: Histograms of expected value distances  $d_E(O_{a_S|J}^{a_z}, O_{Exp}^{a_z})$  (a), and of the geometrical distances  $d_G(O_{a_S|J}^{a_z}, O_{Exp}^{a_z})$  (b). In red the fitting of the histogram with the gamma function (calculated with the Matlab function `histfit`).

have been carried on (i.e. 1250 explorations were considered in this experiment).

In Section 7.1 we make a qualitative analysis of the distributions of the distances computed using the two metrics discussed in Definitions 12 and 13. Section 7.2 discusses the results of the Wilcoxon signed-rank test on the measure of distances. Finally, Sect. 7.3 provides a qualitative analysis of the dynamics of the results varying the two parameters of the experiment, namely the probability of connections  $P^L$  and the bootstrap time  $\#B$ .

### 7.1 Distributions of Distances

We first analysed the distribution of the distances between the ground truth and when using Jøsang’s operators (Figures 2a and 2b respectively);  $\circ_1$  and  $\Gamma_1$  (Figures 3a and 3b);  $\circ_2$  and  $\Gamma_1$  (Figures 3c and 3d),  $\circ_3$  and  $\Gamma_1$  (Figures 3e and 3f);  $\circ_n$  and  $\Gamma_1$  (Figures 3g and 3h).

Visual inspection of Figures 2a, 3a, 3c, 3e, 3g indicates that the distances computed using the expected value distance (Definition 13) can be approximated (qualitatively<sup>8</sup>) using a gamma function, regardless of the choice of the operator.

This result looks reasonable with respect to the opinions computed using Jøsang’s operators (Fig. 2a) due to the fact that the experiment considered the beta reputation system (Ismail and Jøsang 2002) and due to the statistical properties of Jøsang’s operators (Jøsang et al. 2006). It is, however, interesting to note that the use of  $\circ_1$  and  $\Gamma_1$ , or  $\circ_2$  and  $\Gamma_1$ , or  $\circ_3$  and  $\Gamma_1$ , or  $\circ_n$  and  $\Gamma_1$ , all result in similar graphs.

More interesting is the fact that, considering the graphical distance (Figures 2b, 3b, 3d, 3f, 3h) we can conclude that (qualitatively) Jøsang’s operators (and similarly  $\circ_3$  with  $\Gamma_1$ ) are computing opinions whose geometric distance from the ground truth is not distributed using a gamma function

(Figures 2b and 3f). On the other hand, using either  $\circ_1$  with  $\Gamma_1$  (Fig. 3b), or  $\circ_2$  with  $\Gamma_1$  (Fig. 3d), or  $\circ_n$  with  $\Gamma_1$  (Fig. 3h) returns opinions whose geometric distance from the ground truth has an interesting and regular shape, which is similar to a gamma function or a Lognormal distribution. A comprehensive study of this is beyond the scope of the present paper and is left for future work.

### 7.2 Analysis Using the Wilcoxon Test

Since the Kolmogorov–Smirnov tests reported that the distribution of the differences between each pair of distance  $\langle d_E(O_{a_S|\circ}^{a_z}, O_{Exp}^{a_z}), d_E(O_{a_S|J}^{a_z}, O_{Exp}^{a_z}) \rangle$  (resp.  $\langle d_G(O_{a_S|\circ}^{a_z}, O_{Exp}^{a_z}), d_G(O_{a_S|J}^{a_z}, O_{Exp}^{a_z}) \rangle$ ) are significantly different from a normal distributions ( $p < 0.001$ ), we analysed them using the Wilcoxon Signed-Rank Test (WSRT).

This test allows us to conclude whether or not the median of the differences of such pairs of distances is statistically equal to 0. Moreover, looking at the median of the distribution of each distance, we can also verify the significance of the direction of the difference. In other words, if the distribution of  $d_E(O_{a_S|\circ}^{a_z}, O_{Exp}^{a_z})$  has median equal to  $a$  and the median of the distribution of  $d_E(O_{a_S|J}^{a_z}, O_{Exp}^{a_z})$  is  $b$ , we can verify the hypothesis that the difference is significantly positive (if  $a > b$ ) or negative (if  $a < b$ ). Therefore, if the difference is significant, this test shows that one distance from the ground truth is significantly higher than the other. Furthermore, WSRT calculates the sum of the ranks of the pairwise positive differences  $s^+$  and negative differences  $s^-$ . This can be used to indicate the size of this difference: we consider the following simple formula for determining this size, which turns out to be our measure of increment of performance, namely  $(s^+ - s^-)/(s^+ + s^-)$ .

For improving the readability of the results, we grouped the results of the WSRT test according to the type of distance used. Table 2 shows the results of the WSRT considering measure computed using the expected value distance

<sup>8</sup> Figures 2a, 2b, 3a, 3b, 3c, 3d, 3e, 3f, 3g, 3h, are obtained using the Matlab function `histfit`.

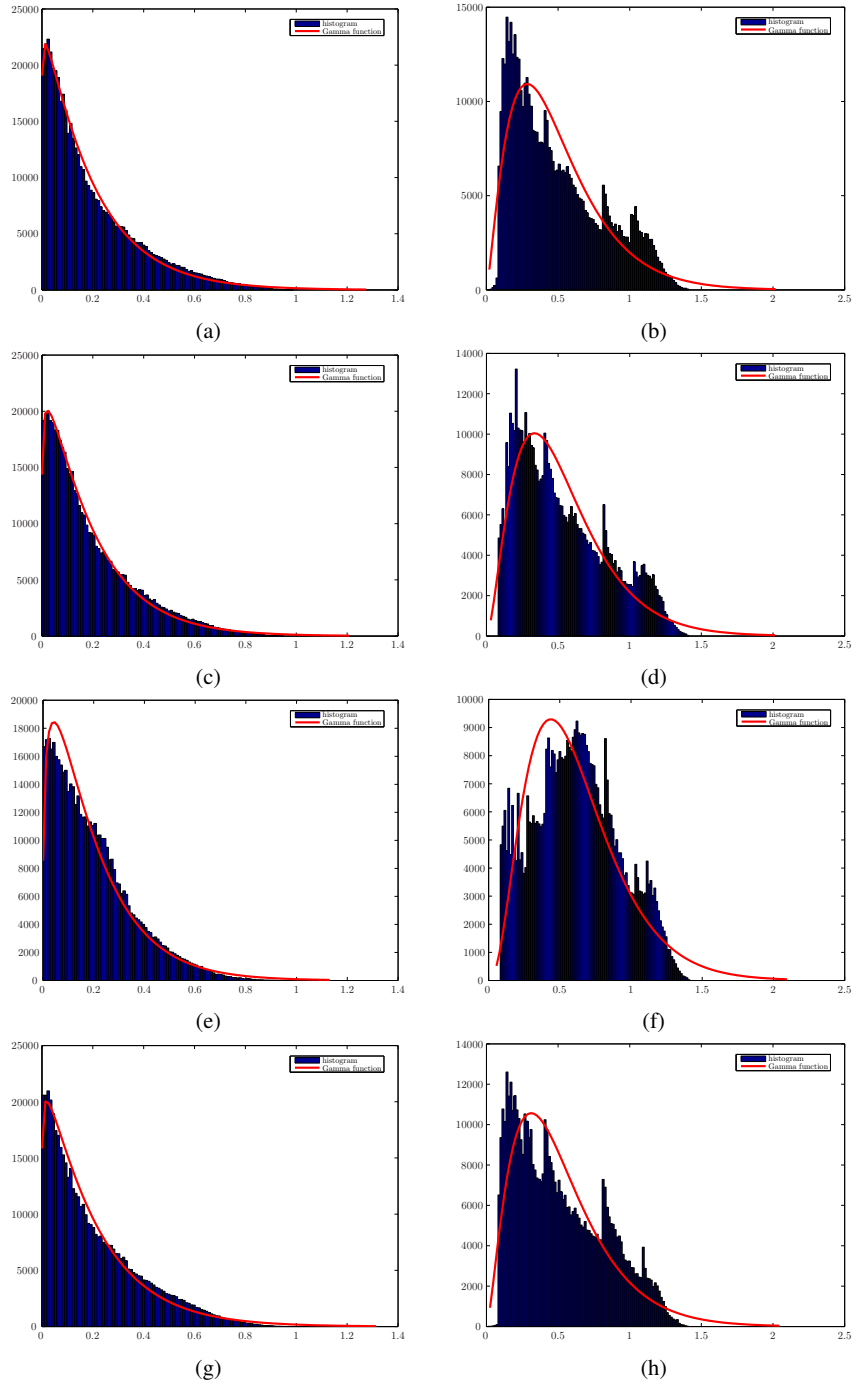


Fig. 3: Histograms of expected value distances  $d_E(O_{a_S|o}^{a_z}, O_{Exp}^{a_z})$  (resp. geometrical distances  $d_G(O_{a_S|o}^{a_z}, O_{Exp}^{a_z})$ ) using the graphical discount operators of Defs. 7 and 10 — i.e.  $\circ = \circ_1$  (a) (resp. (b)),  $\circ = \circ_2$  (c) (resp. (d)),  $\circ = \circ_3$  (e) (resp. (f)),  $\circ = \circ_n$  (g) (resp. (h)). In red the fitting of the histogram with the gamma function (calculated with the Matlab function `histfit`).

Operator	Md* $d_E(O_{a_S o}^{a_z}, O_{Exp}^{a_z})$	Md* $d_E(O_{a_S J}^{a_z}, O_{Exp}^{a_z})$	$s^- (\times 10^{10})$	$s^+ (\times 10^{10})$	$z$	Incr. Performance <sup>†</sup>
$\circ_1$	0.141	0.144	4.11	4.53	-27.457 <sup>‡</sup>	$\approx +5\%$
$\circ_n$	0.156	0.155	4.40	3.95	-29.586 <sup>‡</sup>	$\approx -5\%$
$\circ_2$	0.143	0.142	4.58	3.89	-45.559 <sup>‡</sup>	$\approx -8\%$
$\circ_3$	0.163	0.145	5.12	3.51	-104.098 <sup>‡</sup>	$\approx -19\%$

\* Median; <sup>†</sup> computed as  $(s^+ - s^-)/(s^+ + s^-)$ ; <sup>‡</sup> $p < 0.001$ .

Table 2: Wilcoxon signed-rank significance tests of distances derived using the expected value distance  $d_E(\cdot, \cdot)$ . In grey are the cases where the proposed operator  $\circ \in \{\circ_1, \circ_2, \circ_3, \circ_n\}$  did not outperform Jøsang’s operators. Results are ordered by the increment of performance in descending order (i.e. the first row is the best one).

Operator	Md* $d_G(O_{a_S o}^{a_z}, O_{Exp}^{a_z})$	Md* $d_G(O_{a_S J}^{a_z}, O_{Exp}^{a_z})$	$s^- (\times 10^{10})$	$s^+ (\times 10^{10})$	$z$	Incr. Performance <sup>†</sup>
$\circ_1$	0.415	0.585	1.91	6.70	-310.462 <sup>‡</sup>	$\approx +56\%$
$\circ_n$	0.454	0.608	1.93	6.42	-297.432 <sup>‡</sup>	$\approx +54\%$
$\circ_2$	0.457	0.584	2.31	6.08	-248.968 <sup>‡</sup>	$\approx +45\%$
$\circ_3$	0.607	0.593	4.92	3.72	-77.760 <sup>‡</sup>	$\approx -14\%$

\* Median; <sup>†</sup> computed as  $(s^+ - s^-)/(s^+ + s^-)$ ; <sup>‡</sup> $p < 0.001$ .

Table 3: Wilcoxon signed-rank significance tests of distances derived using the geometric distance  $d_G(\cdot, \cdot)$ . In grey are the cases where the proposed operator  $\circ \in \{\circ_1, \circ_2, \circ_3, \circ_n\}$  did not outperform Jøsang’s operators. Results are ordered by the increment of performance in descending order (i.e. the first row is the best one).

$d_E(\cdot, \cdot)$ , while Table 3 shows the results of the WSRT considering measure computed using the geometrical distance  $d_G(\cdot, \cdot)$ .

From Tables 2 and 3 we can conclude that, regardless the choice of the operator and the type of measure used, the difference between the opinions determined with the proposed operators and the Jøsang’s is significantly ( $p < 0.001$ ).

Concerning the expected value distances, from Table 2 we can see that the WSRT highlights that in the case that  $\circ_1$  with  $\Gamma_1$  is used, the derived opinion is significantly ( $\approx +5\%$ ) closer to the ground truth than the opinion computed using Jøsang’s operators. This is not true according to the other choices of operators, which return opinions whose distance from the ground truth is greater (between 5% and 19%) than Jøsang’s operator.

However, if we consider the graphical distances, from Table 3 we can see that (in order)  $\circ_1$ ,  $\circ_n$ , and  $\circ_2$ , each of which with  $\Gamma_1$ , outperform Jøsang’s operators. Comparing these increments of performance, we can also see that the opinions derived using these operators are much closer to the ground truth (between  $\approx +56\%$  and  $\approx +45\%$ ) than Jøsang’s operators.

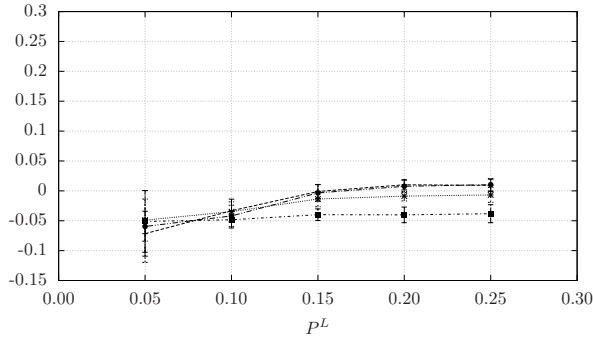
### 7.3 Results w.r.t. Experiment Parameters

Considering the dynamics of the results, Figure 4 depicts the mean and the standard deviation<sup>9</sup> of  $r_E(a_z)$  and  $r_G(a_z)$  for each set of operators used — viz.  $\circ_1$  and  $\Gamma_1$ ,  $\circ_2$  and  $\Gamma_1$ ,  $\circ_n$  and  $\Gamma_1$  — w.r.t. the two variables considered, namely the probability of connections  $P^L$  (Figures 4a and 4b), and the bootstrap time  $\#_B$  (Figures 4c and 4d).

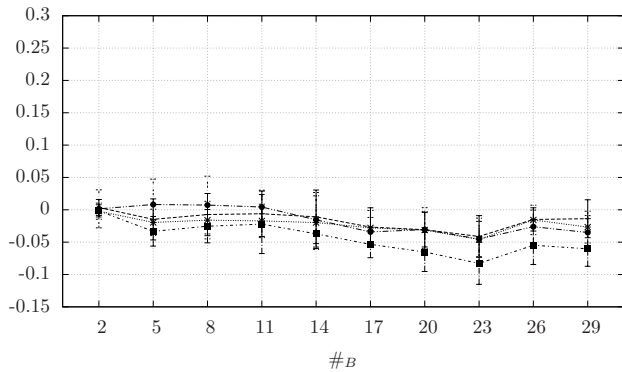
Considering the distances computed using the expected value distance measure, from Figure 4a we can infer that on average Jøsang’s operators are performing better for small values of probability of connections  $P^L$ , and the greater the  $P^L$ , the better are the performance of operators  $\circ_1$ ,  $\circ_n$  and  $\circ_2$  (each of which with  $\Gamma_1$ ). A visual inspection of Figure 4c, however, does not highlight any specific pattern or regularity in the dynamics of the system varying  $\#_B$  (considering expected value distance).

On the other hand, if we consider the results derived using the geometric distance, Fig. 4b qualitatively shows that the greater the probability of connections  $P^L$ , the more similar the operators we propose in this paper are to Jøsang’s. In fact, the more connected the network, the more the bootstrapping phase is important, and this is independent of the

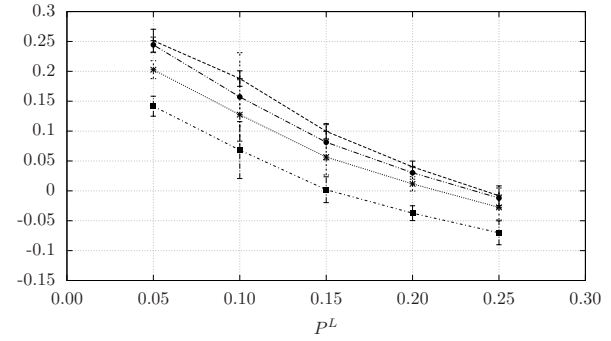
<sup>9</sup> In Sects. 7.1 and 7.2 we show that the distances are not normally distributed and thus from a statistical point of view medians rather than means should be considered. Here, however, we are more interested in the qualitative dynamics of values obtained by varying the parameters of the experiment, and thus we rely on graphical representation of mean and standard deviation.

Mean and Standard Deviation of  $\overline{r_E(a_z)}$  w.r.t. the probability of connections  $P^L$ 

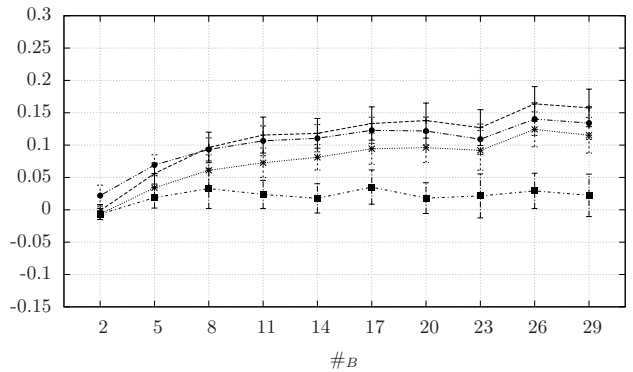
(a)

Mean and Standard Deviation of  $\overline{r_E(a_z)}$  w.r.t. the bootstrap time  $\#_B$ 

(c)

Mean and Standard Deviation of  $\overline{r_G(a_z)}$  w.r.t. the probability of connections  $P^L$ 

(b)

Mean and Standard Deviation of  $\overline{r_G(a_z)}$  w.r.t. the bootstrap time  $\#_B$ 

(d)

Fig. 4: Mean and Standard Deviation of the scalar comparison values using distances of expected values  $\overline{r_E(a_z)}$  (resp. using graphical distances  $\overline{r_G(a_z)}$ ) w.r.t.  $P^L$  (a) (resp. (b)), and w.r.t.  $\#_B$  (c) (resp. (d)). The scale is logarithmic.

choice of operators. However, when we are considering the dynamics of the bootstrapping phase (Fig. 4d), we conclude that the smaller the uncertainty (i.e. the greater the number of interactions among the agents during the bootstrapping phase), the better the proposed operators perform. It is worth to notice that for  $\#_B = 2$ , which leads to a high uncertain opinions,  $\circ_1$  and  $\Gamma_1$ ,  $\circ_2$  and  $\Gamma_1$ ,  $\circ_3$  and  $\Gamma_1$  perform similarly to Jøsang's operators, while choosing  $\circ_n$  and  $\Gamma_1$  leads to a significantly better result. We will investigate this interesting results in future work.

#### 7.4 Summary

To summarise our empirical evaluation, we observe that:

1. the operators  $\circ_1, \circ_2, \circ_3, \circ_n$  (in conjunction to  $\Gamma_1$ ), similarly to Jøsang's operators, return opinions whose expected value distance distribution from the ground truth is close to a gamma function (Figures 2a, 3a, 3c, 3g);
2. the operators  $\circ_1, \circ_2, \circ_n$  (in conjunction with  $\Gamma_1$ ), differ from Jøsang's operators and  $\circ_3$  with  $\Gamma_1$ , and return opinions whose geometric distance distribution from the ground truth shows some qualitative regularity resembling a gamma function or a Lognormal distribution (Figures 2b, 3b, 3d, 3h);
3. the operator  $\circ_1$  with  $\Gamma_1$  outperforms Jøsang's operators in a statistically significant manner, both considering the expected value distance ( $\approx +5\%$ ) and the geometrical distance ( $\approx +56\%$ );
4. the rank of operators (each of which used in conjunction with  $\Gamma_1$ ) w.r.t. their performance is independent from the choice of expected value distance, or geometrical distance, and is as follows:  $\circ_1 \succ \circ_n \succ \circ_2 \succ \circ_3$ ;
5. the less the probability of connections, the more  $\circ \in \{\circ_1, \circ_2, \circ_3, \circ_n\}$  returns opinions closer (according to the graphical distance) to the ground truth than Jøsang's operators;

6. the less the uncertainty (i.e. the more the bootstrap time), the more  $\circ \in \{\circ_1, \circ_2, \circ_3, \circ_n\}$  returns opinions closer (according to the graphical distance) to the ground truth than Jøsang's operators.

## 8 Conclusions and Future Works

The discount and the fusion operators play an important role in standard Subjective Logic, and form the core of the beta reputation system. In fact, they are used to combine and discount reputation information from multiple agents within a trust network.

In this paper, following our earlier work (Cerutti et al. 2013a,b), we introduced a set of intuitive desiderata that operators for discounting and fusion of opinions should provide. From these, we derived a set of requirements and a family of operators, and proved that these satisfy the desiderata, while Jøsang's operators do not. We empirically evaluated<sup>10</sup> the derived operators in a trust scenario and the results shown in Section 7 suggest that:

- one operator taken from the family satisfying the desiderata always outperforms Jøsang's operators;
- according to the geometrical distance among opinions, most of the operators satisfying the desiderata outperform Jøsang's operators;
- there are relationships between the structure of the trust network and the achieved increments of performance.

In particular, the Wilcoxon signed-rank significance test discussed in Section 7.2, shows that the discounting operator ( $\circ_1$ ), used in conjunction with the fusion operator  $\Gamma_1$ , returns opinions closer to the ground truth than Jøsang's operators of 5% considering the expected value distance, and of 56% considering the graphical distance. Therefore, it seems that allowing a reduction of the amount of uncertainty in discounting opinion results on an increment of the performance not only geometrically, but also when the expected values are considered.

An empirical evaluation of the graphical operators on real cases, e.g. (Guha et al. 2004), is already envisaged as the main future work. For example, in a situation where several sensors monitor the same phenomenon (i.e. different intrusion detection systems on the same network), the experiment described in Section 6 can be adapted for assessing the probability of false negatives according to the trustworthiness of each system. The result would be a portfolio of sensors (i.e. a “super” intrusion detection system) able to automatically leverage the strength of each member of the portfolio.

As a side effect, this will provide useful insights about our assumptions regarding the probability of an agent to

“lie” — or, in this case, to provide a false positive/negative results. In this paper we assumed a uniform distribution, but as part of future work we want to investigate the effects of adopting different distributions. In addition, we want to develop graphical operators analogous to other Subjective Logic operators, and we intend to study these, as well as investigate their properties.

## References

- Bisdikian C, Sensoy M, Norman TJ, Srivastaval MB (2012) Trust and obfuscation principles for quality of information in emerging pervasive environments. In: Proceedings of International Workshop on Information Quality and Quality of Service for Pervasive Computing (IQ2S), pp 44–49
- Burnett C, Oren N (2012) Sub-delegation and trust. In: Proceedings of the 11th International Conference on Autonomous Agents and Multiagent Systems, pp 1359–1360
- Castelfranchi C, Falcone R (2010) Trust theory: A socio-cognitive and computational model. Wiley Series in Agent Technology
- Cerutti F, Toniolo A, Oren N, Norman T (2013a) An empirical evaluation of geometric subjective logic operators. In: Chesñevar CI, Onaindia E, Ossowski S, Vouras G (eds) Proceedings of the Second Conference on Agreement Technologies, Lecture Notes in Computer Science, vol 8068, Springer, pp 90–104
- Cerutti F, Toniolo A, Oren N, Norman TJ (2013b) Context-dependent trust decisions with subjective logic. Tech. Rep. ABDN-CS-2013-05-01, Department of Computing Science, University of Aberdeen, URL <http://homepages.abdn.ac.uk/f.cerutti/pages/papers/techreps/ABDN-CS-2013-05-01.pdf>
- Dempster AP (1968) A Generalization of Bayesian Inference. *Journal of the Royal Statistical Society Series B (Methodological)* 30(2):pp. 205–247
- Guha R, Kumar R, Raghavan P, Tomkins A (2004) Propagation of trust and distrust. In: Proceedings of the 13th conference on World Wide Web, ACM Press, pp 403–412
- Haghpanah Y, des Jardins M (2012) Prep: a probabilistic reputation model for biased societies. In: Proceedings of the 11th International Conference on Autonomous Agents and Multiagent Systems, pp 315–322
- Ismail R, Jøsang A (2002) The beta reputation system. In: Proceedings of the 15th Bled Electronic Commerce Conference, pp 324–337
- Jøsang A (2001) A logic for uncertain probabilities. *International Journal of Uncertainty, Fuzziness and Knowledge-Based Systems* 9(3):279–311
- Jøsang A (2008) Cumulative and averaging unfusion of beliefs. In: Proceedings of the International Conference on Information Processing and Management of Uncertainty
- Jøsang A, McAnally D (2004) Multiplication and comultiplication of beliefs. *International Journal of Approximate Reasoning* 38(1):19–51
- Jøsang A, Pope S, Daniel M (2005) Conditional deduction under uncertainty. In: Proceedings of the 8th European Conference on Symbolic and Quantitative Approaches to Reasoning with Uncertainty, pp 824–835
- Jøsang A, Marsh S, Pope S (2006) Exploring different types of trust propagation. In: Stølen K, Winsborough WH, Martinelli F, Masciaci F (eds) Trust Management, Springer, pp 179–192
- Jøsang A, Ismail R, Boyd C (2007) A survey of trust and reputation systems for online service provision. *Decision Support Systems* 43(2):618–644

<sup>10</sup> The code used for this experiment is available at <https://sourceforge.net/projects/slef/>.



- Jøsang A, Azderska T, Marsh S (2012) Trust transitivity and conditional belief reasoning. In: Dimitrakos T, Moona R, Patel D, McKnight DH (eds) Proceedings of 6th IFIPTM International Conference on Trust Management, Springer, pp 68–83
- McAnally D, Jøsang A (2004) Addition and subtraction of beliefs. In: Proceedings of the conference on Information Processing and Management of Uncertainty in Knowledge-Based Systems
- Sabater J, Sierra C (2005) Review on computational trust and reputation models. Artificial Intelligence Review 24:33–60
- Sensoy M, Fokoue A, Pan JZ, Norman TJ, Tang Y, Oren N, Sycara K (2013) Reasoning about uncertain information and conflict resolution through trust revision. In: Proceedings of the 2013 international conference on Autonomous agents and multi-agent systems
- Shafer G (1976) A Mathematical Theory of Evidence. Princeton University Press
- Urbano J, Rocha A, Oliveira E (2013) A socio-cognitive perspective of trust. In: Ossowski S (ed) Agreement Technologies, Law, Governance and Technology Series, vol 8, Springer, pp 419–429

## A The Geometry of Subjective Logic

A SL opinion  $O \triangleq \langle b_O, d_O, u_O \rangle$  is a point in the  $\mathbb{R}^3$  space, identified by the coordinate  $b_O$  for the first axis,  $d_O$  for the second axis, and  $u_O$  for the third axis. However, due to the requirement that  $b_O + d_O + u_O = 1$ , an opinion is a point inside (or at least on the edges of) the triangle  $\triangle BDU$  shown in Fig. 5, where  $B = \langle 1, 0, 0 \rangle$ ,  $D = \langle 0, 1, 0 \rangle$ ,  $U = \langle 0, 0, 1 \rangle$ .

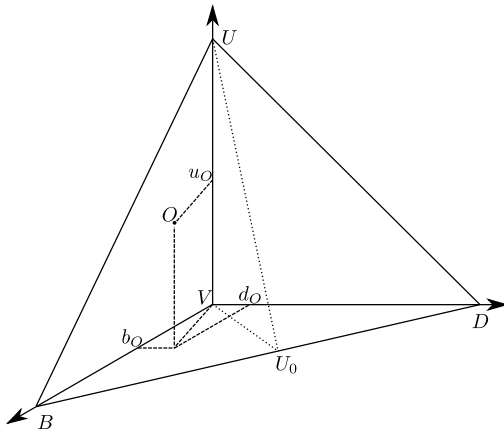


Fig. 5: The Subjective Logic plane region.

**Definition 14** The *Subjective Logic plane region*  $\triangle BDU$  is the triangle whose vertices are the points  $B \triangleq \langle 1, 0, 0 \rangle$ ,  $D \triangleq \langle 0, 1, 0 \rangle$ , and  $U \triangleq \langle 0, 0, 1 \rangle$  on a  $\mathbb{R}^3$  space where the axes are respectively the one of belief, disbelief, and uncertainty predicted by SL.

Since an opinion is a point inside triangle  $\triangle BDU$ , it can be mapped to a point in Fig. 6. This representation is similar to the one used in (Jøsang 2001) for representing opinions in SL, but here the belief and disbelief axes are swapped.

In order to keep the discussion consistent with Jøsang's work (Jøsang 2001), in what follows we will scale triangle  $\triangle BDU$  by a factor  $1 : \frac{\sqrt{3}}{\sqrt{2}}$  thus obtaining that  $|\vec{B_0B}| = |\vec{D_0D}| = |\vec{U_0U}| = 1$ .

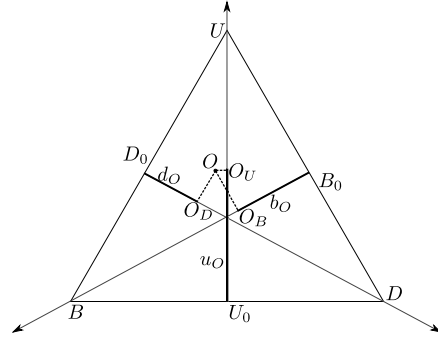


Fig. 6: An opinion  $O \triangleq \langle b_O, d_O, u_O \rangle$  in SL after the  $1 : \frac{\sqrt{3}}{\sqrt{2}}$  scale. The belief axis is the line from  $B_0$  (its origin) toward the  $B$  vertex, the disbelief axis is the line from  $D_0$  toward the  $D$  vertex, and the uncertainty axis is the line from  $U_0$  toward the  $U$  vertex.

These geometric relations lie at the heart of the Cartesian transformation operator, which is the subject of the next subsection.

### A.1 The Cartesian Representation of Opinions

As shown in A, an opinion in SL can be represented as a point in a planar figure (Fig. 6) laying on a Cartesian plane. In this section we will introduce the Cartesian transformation operator, which returns the Cartesian coordinate of an opinion.

First of all, let us define the axes of the Cartesian system we will adopt.

**Definition 15** Given the SL plane region  $\triangle BDU$ , the *associated Cartesian system* is composed by two axes, named respectively  $x, y$ , where the unit vector of the  $x$  axis  $\vec{e}_x = \frac{1}{|\vec{BD}|} \vec{BD}$ , the unit vector of the  $y$  axis  $\vec{e}_y = \vec{e}_u$ , and  $B$  is the origin.

Figure 7 depicts this Cartesian system.

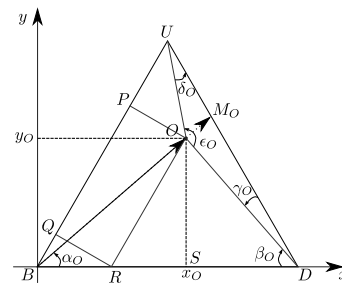


Fig. 7: An opinion and its representation in the Cartesian system.

The correspondence between the three values of an opinion and the corresponding coordinate in the Cartesian system we defined is shown in the following proposition.

**Proposition 4** (Cerutti et al. 2013b, Prop. 1) *Given a SL plane region  $\triangle BDU$  and its associated Cartesian system  $\langle x, y \rangle$ , an opinion  $O \triangleq \langle b_O, d_O, u_O \rangle$  is identified by the coordinate  $\langle x_O, y_O \rangle$  s.t.:*

$$\begin{aligned}
- x_O &\triangleq \frac{d_O + u_O \cos(\frac{\pi}{3})}{\sin(\frac{\pi}{3})} \\
- y_O &\triangleq u_O
\end{aligned}$$

*Proof* Proving that  $y_O \triangleq u_O$  is trivial.

Let us focus on the first part of the proposition. Consider Figure 7. Given  $O$ , we note that the for the point  $P$ ,  $\frac{1}{|\overrightarrow{PO}|} \overrightarrow{PO} = \vec{e}_b$  (i.e.  $\overrightarrow{PO}$  is parallel to the disbelief axis) and  $\frac{1}{|\overrightarrow{BP}|} \overrightarrow{BP} = \frac{1}{|\overrightarrow{BU}|} \overrightarrow{BU}$  (i.e.  $P$  is on the line  $\overrightarrow{BU}$ ), and therefore  $\angle_{BPO} = \frac{\pi}{2}$ . Then we must determine  $Q$  and  $R$  s.t.  $\overrightarrow{QR} = \overrightarrow{PO}$  and  $y_R = 0$ . By construction  $|\overrightarrow{PO}| = |\overrightarrow{QR}| = d_O$ ,  $\angle_{QRB} = \frac{\pi}{6}$ ,  $\angle_{ORD} = \frac{\pi}{3}$ , and  $x_O \triangleq |\overrightarrow{BS}| = |\overrightarrow{BR}| + |\overrightarrow{RS}|$ , where  $|\overrightarrow{BR}| = \frac{d_O}{\sin(\frac{\pi}{3})}$ , and  $|\overrightarrow{RS}| = \frac{u_O}{\sin(\frac{\pi}{3})} \cos(\frac{\pi}{3})$ .

There are some notable elements of Fig. 7 that we will repeatedly use below, and we therefore define them as follows:

- the angle  $\alpha_O$  determined by the  $x$  axis and the vector  $\overrightarrow{BO}$ ;
- the three angles ( $\gamma_O, \delta_O$ , and  $\epsilon_O$ ) of the triangle  $ODU$ , namely the triangle determined by linking the point  $O$  with the vertex  $D$  and  $U$  through straight lines.

**Definition 16** Given the SL plane region  $\triangle BDU$ , given  $O = \langle b_O, d_O, u_O \rangle$  whose coordinates are  $\langle x_O, y_O \rangle$  where  $x_O \triangleq \frac{d_O + u_O \cos(\frac{\pi}{3})}{\sin(\frac{\pi}{3})}$

and  $y_O \triangleq u_O$ , let us define and (via trivial trigonometric relations) compute the following.

$$\begin{aligned}
- \alpha_O &\triangleq \angle_{OBD} = \begin{cases} 0 & \text{if } b_O = 1 \\ \arctan\left(\frac{u_O \sin(\frac{\pi}{3})}{d_O + u_O \cos(\frac{\pi}{3})}\right) & \text{otherwise;} \end{cases} \\
- \beta_O &\triangleq \angle_{ODB} = \begin{cases} \frac{\pi}{3} & \text{if } d_O = 1 \\ \arctan\left(\frac{u_O \sin(\frac{\pi}{3})}{1 - (d_O + u_O \cos(\frac{\pi}{3}))}\right) & \text{otherwise;} \end{cases} \\
- \gamma_O &\triangleq \angle_{ODU} = \frac{\pi}{3} - \beta_O; \\
- \delta_O &\triangleq \angle_{OUD} = \begin{cases} 0 & \text{if } u_O = 1 \\ \arcsin\left(\frac{b_O}{|\overrightarrow{OU}|}\right) & \text{otherwise;} \end{cases} \\
- \epsilon_O &\triangleq \angle_{DOU} = \pi - \gamma_O - \delta_O;
\end{aligned}$$

where  $|\overrightarrow{OU}| = \sqrt{\frac{1}{3}(1 + d_O - u_O)^2 + b_O^2}$ .

The angle  $\alpha_O$  is called the *direction* of  $O$ .

Equivalently, we can write  $\overrightarrow{BO}$  or  $\langle B, \alpha_O, |\overrightarrow{BO}| \rangle$ .

Finally, as an element of SL is bounded to have its three components between 0 and 1, we are also interested in determining the point  $M_O$  such that the vector  $\overrightarrow{BM_O}$  has the maximum magnitude given (a) the direction  $\alpha_O$  of an opinion  $O$ , and (b)  $M_O$  is a SL opinion. In other words, determining the magnitude of  $\overrightarrow{BM_O}$  will allow us to re-define the vector  $\overrightarrow{BO}$  as a fraction of  $\overrightarrow{BM_O}$ .

**Definition 17** Given the SL plane region  $\triangle BDU$ , and  $O \triangleq \langle b_O, d_O, u_O \rangle$  whose coordinates are  $\langle x_O, y_O \rangle$  where  $x_O \triangleq \frac{d_O + u_O \cos(\frac{\pi}{3})}{\sin(\frac{\pi}{3})}$  and  $y_O \triangleq u_O$ , and  $\alpha_O \triangleq \angle_{OBD} = \arctan\left(\frac{u_O \sin(\frac{\pi}{3})}{d_O + u_O \cos(\frac{\pi}{3})}\right)$ ,

let us define  $M_O \triangleq \langle x_{M_O}, y_{M_O} \rangle$  as the intersection of the straight line passing for  $O$  and  $B$ , and the straight line passing for  $U$  and  $D$ , and thus define the following.

$$\begin{aligned}
- x_{M_O} &\triangleq \frac{2 - y_O + \tan(\alpha_O) x_O}{\tan(\alpha_O) + \sqrt{3}}; \\
- y_{M_O} &\triangleq -\sqrt{3} x_{M_O} + 2.
\end{aligned}$$

## B Proofs

**Proposition 1** Given the two opinions  $T = \langle b_T, d_T, u_T \rangle$  and  $C = \langle b_C, d_C, u_C \rangle$ , and  $W = T \circ_n C$ , the naïve-discount of  $C$  by  $T$ , then:

- i.  $W \in \mathbb{O}$ ;
- ii. if  $C = \langle 1, 0, 0 \rangle$ , then  $W = T$ , i.e. requirement **rd1**;
- iii. if  $C = \langle 0, 0, 1 \rangle$ , then  $W = C$ , i.e. requirement **rd2**;
- iv.  $b_W \leq b_T$ , i.e. requirement **rd3**.

*Proof*

- i. Let us prove that  $0 \leq b_W \leq 1, 0 \leq d_W \leq 1, 0 \leq u_W \leq 1$ .

To prove that  $b_W \geq 0, d_W \geq 0$ , and  $u_W \geq 0$  is trivial since  $C$  and  $T$  are opinions.

$b_W = b_C \cdot b_T \leq 1$  is immediate since  $C$  and  $T$  are opinions.

$d_W = b \cdot d_T + d_C \leq 1$  can be rewritten as  $d_T \leq 1 + \frac{u_C}{b_C}$  if  $b_C \neq 0$ , or  $d_C \leq 1$  otherwise. Both in-equations are verified since  $C$  and  $T$  are opinions.

$u_W = b \cdot u_T + u_C \leq 1$  can be rewritten as  $d_T \leq 1 + \frac{d_C}{b_C}$  if  $b_C \neq 0$ , or  $u_C \leq 1$  otherwise. Both in-equations are verified since  $C$  and  $T$  are opinions.

Finally,  $b_W + d_W + u_W = b_C(b_T + d_T + u_T) + d_C + u_C = 1$

- ii. Given  $C = \langle 1, 0, 0 \rangle$ ,  $W = T \circ_n C$  is such that:

$$\begin{aligned}
- b_W &= b_C \cdot b_T = b_T; \\
- d_W &= b_C \cdot d_T + d_C = d_T; \\
- u_W &= b_C \cdot u_T + u_C = u_T.
\end{aligned}$$

- iii. Given  $C = \langle 0, 0, 1 \rangle$ ,  $W = T \circ_n C$  is such that:

$$\begin{aligned}
- b_W &= b_C \cdot b_T = 0 = b_C; \\
- d_W &= b_C \cdot d_T + d_C = d_C; \\
- u_W &= b_C \cdot u_T + u_C = u_C.
\end{aligned}$$

- iv. By contradiction,  $b_W = b_C \cdot b_T > b_T$  leads to  $b_C > 1$ , which is impossible.  $\square$

**Proposition 2** Given an opinion  $T = \langle b_T, d_T, u_T \rangle$ , and its Cartesian representation, the four-sided figure  $PQDU$  represents  $\mathbb{O}_T$ , where:

- $P \triangleq \langle b_T, 0, 1 - b_T \rangle$ ; and
- $Q \triangleq \langle b_T, 1 - b_T, 0 \rangle$ .

*Proof* By Definition 8,  $\mathbb{O}_T = \{X \in \mathbb{O} | b_X \leq b_T\}$ . From Prop. 4,

$$\begin{aligned}
b_X &\leq b_T \\
y_X &\geq \sqrt{3}x_X + 2(1 - b_T) \\
u_X &\geq 1 - d_X - b_T
\end{aligned}$$

Therefore:

- if  $d_X = 0, u_X \geq 1 - b_T$  (limit case  $\langle b_T, 0, 1 - b_T \rangle$ );
- if  $u_X = 0, d_X \geq 1 - b_T$  (limit case  $\langle b_T, 1 - b_T, 0 \rangle = Q$ ).  $\square$

**Theorem 1** Given  $T = \langle b_T, d_T, u_T \rangle$ ,  $C = \langle b_C, d_C, u_C \rangle$ , and  $W = T \circ_{\alpha_{C'}} C$ , then:

- i.  $W = \langle b_W, d_W, u_W \rangle$  is an opinion;
- ii. if  $C = \langle 1, 0, 0 \rangle$ , then  $W = T$ ;
- iii. if  $C = \langle 0, 0, 1 \rangle$ , then  $W = C$ ;
- iv.  $b_W \leq b_T$ .

*Proof* Proving the thesis in the limit case is trivial. In the following we will assume, without loss of generality, that  $\alpha_{C'} \neq \frac{\pi}{2}, \alpha_{C'} \neq -\frac{\pi}{3}, \alpha_{C'} \neq \frac{2}{3}\pi$ .

- (i.)  $W = \langle b_W, d_W, u_W \rangle$  must respect

$$u_W + d_W \leq 1 \quad (1)$$

From Def. 9 it is clear that Equation 1 can be rewritten as follows.

$$u_T + d_T + \frac{r_C}{2} \frac{2\sqrt{\tan^2(\alpha_{C'}) + 1}}{|\tan(\alpha_{C'}) + \sqrt{3}|} b_T (\sin(\alpha_{C'}) + \sqrt{3} \cos(\alpha_{C'})) \leq 1 \quad (2)$$

In turn, using the relation  $\tan(\alpha_{C'}) = \frac{\sin(\alpha_{C'})}{\cos(\alpha_{C'})}$ , this can be rewritten as

$$u_T + d_T + r_C b_T \leq 1$$

which entails the requirement that  $r_C \leq \frac{1-u_T-d_T}{b_T} = \frac{b_T}{b_T} = 1$ . However, from definition 9, we know that  $r_C \leq 1$ , fulfilling this requirement.

(ii.)  $C = \langle 1, 0, 0 \rangle$  implies that  $|\vec{BC}| = 0$  and thus  $r_C = 0$ . Therefore, from Def. 9,  $u_W = u_T + \sin(\alpha_{C'}) r_C |\vec{TM_{C'}}| = u_T$  and this result also implies that  $d_W = d_T$ . Since Point 1 shows that  $W$  is an opinion in SL, we conclude that  $W = T$ .

(iii.)  $C = \langle 0, 0, 1 \rangle$  implies  $r_C = 1$ ,  $\frac{\pi}{3} \frac{\epsilon_T}{\pi} - \beta_T \leq \alpha_{C'} \leq \epsilon_T - \beta_T$ , and thus  $\alpha_{C'} = \epsilon_T - \beta_T = \epsilon_T - \beta_T = \frac{2}{3}\pi - \delta_t$ . Therefore, we obtain that  $u_W = u_T + \frac{b_T}{2} (1 + \frac{\sqrt{3}}{\tan(\delta_T)})$ . From Definition 16 and the trigonometric property that  $\tan(\arcsin(v)) = \frac{v}{\sqrt{1-v^2}}$  we obtain that  $u_W = u_T + \frac{b_T}{2} + \frac{\sqrt{3}}{2} \sqrt{|\vec{TU}|^2 - b_T^2}$ . From Definition 16 we can write:

$$u_W = u_T + \frac{b_T}{2} + \frac{\sqrt{3}}{2} \frac{1 + d_T - u_T}{\sqrt{3}} = \frac{1}{2} (1 + b_T + d_T + u_T) = 1 \quad (3)$$

Similarly,  $d_W = d_T + \frac{u_T}{2} - \frac{1}{2} + \frac{\sqrt{3}}{2} \frac{1}{\sin(\delta_T)} b_T = d_T + \frac{u_T}{2} - \frac{1}{2} + \frac{\sqrt{3}}{2} |\vec{TU}|$ . From Def. 16 we have

$$d_W = d_T + \frac{u_T - 1}{2} + \frac{3}{4} b_T - \frac{\sqrt{3}}{4} b_T \frac{1 + d_T - u_T}{\sqrt{3} b_T} = \frac{1}{4} (4d_T + 2u_T - 2 + 3b_T - 1 + u_T - d_T) = 0 \quad (4)$$

From Equations 3 and 4, together with Point 1, it follows that  $W = \langle 0, 0, 1 \rangle = C$ .

(iv.) Suppose instead  $b_W > b_T$ .

$$1 - d_W - u_W > 1 - d_T - u_T$$

$$d_W + u_W < d_T + u_T$$

$$d_T + \sin(\alpha_{C'} + \frac{\pi}{3}) \frac{r_C}{\sin(\alpha_{C'} + \frac{\pi}{3})} + u_T < d_T + u_T$$

$$r_C < 0$$

but  $0 \leq r_C \leq 1$ . *Quod est absurdum.*  $\square$

**Proposition 3** Given the opinions  $T_1, T_2, \dots, T_n$ ,  $C_1, C_2, \dots, C_n$ ,  $W_1, W_2, \dots, W_n$  s.t.  $\forall i \in \{1 \dots n\}$ ,  $W_i = T_i \circ C_i$ , and  $\langle b_{\Gamma_1(W_1, \dots, W_n)}, d_{\Gamma_1(W_1, \dots, W_n)}, u_{\Gamma_1(W_1, \dots, W_n)} \rangle$  the opinion resulting from the fusion of opinions  $W_1, W_2, \dots, W_n$ , then:

i.  $\langle b_{\Gamma_1(W_1, \dots, W_n)}, d_{\Gamma_1(W_1, \dots, W_n)}, u_{\Gamma_1(W_1, \dots, W_n)} \rangle$  is an opinion

$$\text{ii. } x_{\Gamma_1(W_1, \dots, W_n)} = \frac{1}{\sum_{i=1}^n K_i} \left( \sum_{i=1}^n K_i x_{W_i} \right)$$

$$\text{iii. } y_{\Gamma_1(W_1, \dots, W_n)} = \frac{1}{\sum_{i=1}^n K_i} \left( \sum_{i=1}^n K_i y_{W_i} \right)$$

*Proof*

(i.) To prove that  $\langle b_{\Gamma_1(W_1, \dots, W_n)}, d_{\Gamma_1(W_1, \dots, W_n)}, u_{\Gamma_1(W_1, \dots, W_n)} \rangle$  is an opinion, we have to show that  $u_{\Gamma_1(W_1, \dots, W_n)} + d_{\Gamma_1(W_1, \dots, W_n)} \leq 1$  holds.

$$u_{\Gamma_1(W_1, \dots, W_n)} + d_{\Gamma_1(W_1, \dots, W_n)} = \frac{1}{\sum_{i=1}^n K_i} \left( \sum_{i=1}^n K_i (u_{W_i} + d_{W_i}) \right) = \frac{1}{\sum_{i=1}^n K_i} \left( \sum_{i=1}^n K_i (1 - b_{W_i}) \right) = 1 - \frac{1}{\sum_{i=1}^n K_i} \left( \sum_{i=1}^n K_i b_{W_i} \right)$$

(ii.) From Prop. 4,

$$\begin{cases} x_{\Gamma_1(W_1, \dots, W_n)} = \frac{d_{\Gamma_1(W_1, \dots, W_n)}}{\sin(\frac{\pi}{3})} + \frac{1}{2 \sin(\frac{\pi}{3})} \frac{1}{\sum_{i=1}^n K_i} \left( \sum_{i=1}^n K_i u_{W_i} \right) \\ x_{\Gamma_1(W_1, \dots, W_n)} = \frac{1}{\sin(\frac{\pi}{3}) \sum_{i=1}^n K_i} \left( \sum_{i=1}^n K_i (d_{W_i} + \frac{u_{W_i}}{2}) \right) \end{cases}$$

Thus we obtain:

$$\begin{aligned} d_{\Gamma_1(W_1, \dots, W_n)} &= \sin(\frac{\pi}{3}) \left( \frac{1}{\sin(\frac{\pi}{3}) \sum_{i=1}^n K_i} \left( \sum_{i=1}^n K_i (d_{W_i} + \frac{u_{W_i}}{2}) \right) + \right. \\ &\quad \left. - \frac{1}{2 \sin(\frac{\pi}{3}) \sum_{i=1}^n K_i} \left( \sum_{i=1}^n K_i u_{W_i} \right) \right) \\ &= \frac{1}{\sum_{i=1}^n K_i} \left( \left( \sum_{i=1}^n K_i (d_{W_i} + \frac{u_{W_i}}{2}) \right) - \left( \sum_{i=1}^n \frac{u_{W_i}}{2} \right) \right) \\ &= \frac{1}{\sum_{i=1}^n K_i} \left( \sum_{i=1}^n K_i d_{W_i} \right) \end{aligned}$$

Since  $\frac{1}{\sum_{i=1}^n K_i} \left( \sum_{i=1}^n K_i b_{W_i} \right) \geq 0$ , then  $u_{\Gamma_1(W_1, \dots, W_n)} + d_{\Gamma_1(W_1, \dots, W_n)} \leq 1$  holds.

(iii.) Immediate from Prop. 4.  $\square$